7th ISCA Workshop on Speech and Language Technology in Education
25-26 August 2017, Stockholm, Sweden

Emotion and Memory Model for a Robotic Tutor in a Learning Environment

Muneeb Imtiaz Ahmad[1], *Omar Mubin*[2], *Suleman Shahid* [3], *Joanne Orlando*[4]

[1]The MARCS Institute, Human-Machine Interaction Group, Western Sydney University, Australia
[2]School of Computing, Engineering and Mathematics, Western Sydney University, Australia
[3]Department of Computer Science, Lahore University of Management and Sciences, Pakistan
[4]School of Education, Western Sydney University, Australia

muneeb.ahmad@westernsydney.edu.au, o.mubin@westernsydney.edu.au,
suleman.shahid@lums.edu.pk, j.orlando@westernsydney.edu.au

Abstract

In this paper, we present an emotion and memory model for a social robot. The model allowed the robot to create a memory account of a child's emotional events over 4 individual sessions. The robot then adapted its behaviour based on the developed memory. We tested our model through using the NAO robot. The robot was programmed to teach vocabulary to children during the popular game 'Snakes and Ladders'. We conducted a between-subject study with 24 children at a primary school to check the validity of our model. We also evaluated the impact of robot's positive, negative, and neutral emotional feedback of the NAO robot on children vocabulary learning. Three groups of children (8/group) interacted with the robot for four different times during three weeks. Our results showed that the condition where the robot displayed positive emotional responses had a significant effect on the child's learning performance as compared to the two other conditions: negative feedback and neutral feedback. In addition, our model helped children in improving their vocabulary.

Index Terms: Children-Robot Interaction, Education, Long-term Interaction, Adaptive Social Robot.

1. Introduction

One of the growing interests of the social robotics research community in the last decade has been towards the use of social robots in Education. Robots have been employed as a tool to teach computer programming skills in the past [1]. However, due to the introduction of Humanoid robots, new opportunities have opened to use robots in education in ways other than as a tool to teach concepts from various subjects. Most recently, Mubin et al. [2] reported a survey on the applicability of robots in Education. They emphasised the need for an adaptation mechanism that will enable the robot to adapt its behaviour according to the characteristics of the user/student. The applications of such adaptive social robots can be found in education [3] [4]. Most of these studies evaluated short-term interaction [5] and did not capture childrens real interaction with a robot and long term engagement with a robot which is essential to understand the role of robots in the future educational landscape.

We face many technical and social challenges during long-term interaction with social robots [6]. One of them is about decrease in children's interest in the interaction over time. The reasons for this decrease in interest are robot's repetitive behaviour [7] and loss of novelty factor [8]. It is emphasised to implement various autonomous adaptation mechanisms for a social robot to overcome the aforementioned effect. These mechanisms can be based on user's emotions, memory, or personality [5] [9] [8]. The autonomous adaptation mechanism for a social robot can be implemented using different approaches. It can either be through combining machine learning algorithms to chose the robot's behaviour or by following cognitive models that describe how humans create memory or how emotions are regulated in diverse situations and later applying them for social robots. For instance: Belpaeme et al. [10] proposed a model for adaptive strategies for sustainable long-term social interaction based on the theories in cognitive sciences. Trafton et al. [11] presented a cognitive architecture named ACT-R/E (Adaptive Character of Thought-Rational / Embodied) that enables the robot to predict what a user will do in a certain scenario through understanding previous knowledge about the user. Leite et al. [8] designed an emphatic model for an iCAT robot capable of playing chess with children. As stated earlier, we find limited research on social robots that have been used as partners with students in a learning environment during long-term interactions [12]. These afore-listed models are designed for specific purposes in HRI. We therefore, find a vacuum for a robotic model that can be employed to facilitate personalised learning.

In this paper, we present an emotion and memory model for a social robot. The model enables the robot to create a memory of user's emotional events and selects an appropriate behaviour accordingly. The model is based on the theory on how humans create memories of an emotional event/episode [13]. We conducted a 4 week long-term children-robot interaction (cHRI) study to evaluate our model. We programmed the NAO robot to play snakes and ladders game designed to teach vocabulary from the Robot Interaction Language (ROILA) [14] to children in a playful and interactive way. We implemented three types of robot's emotional responses based on positive, negative and neutral emotional events happening during a snakes and ladders game play. Our research question is about studying the effect of robot's positive, negative and neutral emotional response on a child's long-term learning performance in a vocabulary learning task. To the best of our knowledge, this effect has not been studied during the long-term child-robot interaction.

In this study, we choose to focus on vocabulary learning as the interaction task because it is one of the essential components of language learning [15]. Vocabulary learning helps improve listening, reading, writing and speaking skills [16]. Moreover, we choose the artificial language ROILA for this study because it was created based on the rules, syntax, and principles of the major natural languages of the world, which will allow us to mitigate the confounding factor of children having different linguistic backgrounds; it will always be an influence but perhaps less so in the case of ROILA [14] because it has no connection with other languages and dialects spoken. We choose games because the significance of play and interaction in education has

10.21437/SLaTE.2017-4

been well described [17] [18].

2. Emotion and Memory Model

Our model is grounded in the process of formation of emotional memories as described by [13]. It is a well understood fact that humans create memories of both positive and negative emotional experiences. This type of memory is usually stored in two different parts of the human memory system: 1) explicit memory that refers to conscious memory and 2) implicit memory that refers to unconscious memory. In the literature, the memory of the different emotional experiences is categorised as Emotional memory (implicit memory) and memory about emotions (explicit memory). In general, emotional events are processed in human sensory systems. They are later transmitted to the temporal lobe or to the amygdala in order to form either an explicit memory or an implicit memory. Simultaneously, the memory is retrieved in case of occurrence of a cue. The cue is again processed by the sensory system that later leads to retrieval of both explicit or implicit memories. We have utilised the process of creation and retrieval of the memory about the external emotional event as described by [13] and created a model for robots to create memory about different emotional situations and behave accordingly.

It is necessary to define positive and negative emotional events in order to explain our model for the robot. Positive emotional events are described as events when goals are achieved or no immediate problems are encountered towards achieving the goal. Negative emotional events are registered as impediments towards a plan and causing loss to achieve a certain goal. On the other hand, Neutral events are situations that do not significantly threaten an outcome in either positive or negative ways [19].

Levine et al. [20] presented a review on emotion and memory research and showed that different types of information are remembered under various emotional states. As a user's emotional state is directly related to an emotional situation. A positive or negative situation would refer to positive or negative states. It is therefore, necessary to understand the information that should be stored in an emotional state or at an emotional situation. According to [20], humans store a broad range of information from general knowledge and the environment to their memory. Depending on the type of negative emotional state (sad, fearful, or anger) during an emotional situation, humans store different types of information. For example; Sadness may lead to remember about the outcomes and consequences of goal failure. Anger may lead to store information about goals or agents obstructing goal fulfilment. Lastly, fear may lead to storing information about the source of threat and means of avoiding the threat [20].

Based on the general understanding of how humans store emotional information in their memory, we have designed our emotion and memory model, as shown in figure 1, to enable a robot to create a memory of user emotional events and behave accordingly. The purpose of the model is to facilitate different kinds of learning such as concepts from science or mathematics during children-robot-game interactions. Our model has four different modules 1) Inputs, 2) Emotional Event Calculation (EEC), Memory Mechanism Generation (MMG) and 4) Behaviour Selection Unit (BSU). The model has three input types: 1) Game event, 2) user emotional state, and 3) Learning outcome that amalgamates to create an emotional event during the interaction. Based on the type of game event (positive or negative or neutral), user emotional state (happy, sad, angry, fear, surprise, neutral), and learning state, we calculate the type of emotional event in the EEC module. EEC module transmits this information to the MMG module. Based on the type of event and following the type of information remembered under various emotional situations, we send this information to our Memory Processing Unit. We create the memory for the robot in this unit. In addition, in the case of an occurrence of same event type during same circumstances, we update our database with the new event and send the information back to the MMG module. The MMG module later transmits the information to the BSU, that is responsible for selecting an appropriate behaviour or response. Lastly, the robot displays the behaviour.

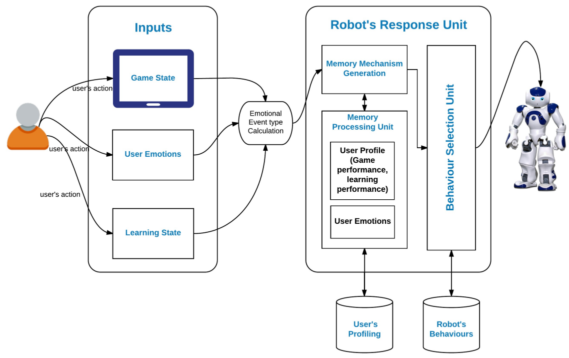

Figure 1: *Emotion and Memory model*

2.1. System Description

To test the applicability of our model, we implemented a scenario where the NAO robot plays snakes and ladder game with a child during an one-to-one interaction. In this section, we present our modified version of the snakes and ladders game. We also discuss the mechanism we used to calculate a type of an emotional event along with the type of information stored in our system. Lastly, we give information on selection of the robot's behaviours under different situations.

Snake and Ladders Game: We modified the game to facilitate vocabulary learning [21]. We updated rules of the Snakes and Ladders game as shown in figure 2 and also introduced stars on the game board. On every snake appearance in the game, NAO robot was programmed to teach a new ROILA word to the child. In the first iteration, the child was asked to go back to the tail of the snake. However, if the same snake reappeared and if the child stated that the word correctly that was taught to him/her on that number, the child could bypass the snake. The same process was repeated on each snake. In the case of a ladder, the child would take the ladder. On every star, a positive or negative number appeared on the dice suggesting the player to move forward or backwards. Lastly, the child was declared the winner when he/she reached the 100 mark.

Figure 2: *Snakes and Ladders*

Applying the Model in the Snakes and Ladders Game: We, in order to realise on the emotional event type, categorised both positive and negative type of game events based on children reactions coded in our previous study on various game events [22]. In the previous study, we coded for the significant game events such as appearance of a snake, ladder or positive/negative star near or away from the 100 mark, continuous sixes on the dice, continuous wastage of turns near 100, and winning or losing the game. We calculated the emotional state of the player through automatic facial scans as described in our previous research [22]. We stored six different emotions (happy, sad, fear, surprise, angry, and neutral) values after every 10 seconds of the interaction. On each significant game event, we calculated current emotional state by taking an average of the last six emotional states stored in our system. Lastly, the learning state refereed to the outcome of the words taught during the interaction. In table 1, we briefly present our list of selected emotional events calculated on the basis of three inputs of our model along with the type of information stored during these events in order to create robot's memory. We also include the behaviour for the NAO robot. For instance, considering the definition of a negative emotional event, a snake near 100 will be rated as a negative event because it hampers the child from winning the game or thwarts the child from achieving the final goal. The information will be stored about the number of times a snake is encountered near 100 and the emotional state of the user. Similarly, a ladder near 100 will be considered a positive event because it is helping the user achieve the end goal. The information about the ladder will be stored in this regard. The behaviour selection of the robot uses the memory of previous emotional events to generate context-aware verbal and nonverbal response either independently or simultaneously. As our purpose was to confirm our model's applicability, we used decision making statements to chose robot's behaviour. We created a database of robot's behaviour consisting of all plausible emotional events during the snakes and ladder game. On each event, the robot displayed the most appropriate behaviour by retrieving it from the database. In the table 1, for understanding, we only enlist a few behaviours.

3. Research Method

Our research tried to explore two aspects. Firstly, we wanted to understand how well our model for the robot performed in terms of teaching vocabulary to children in a long-term interaction. Secondly, we find different claims in literature with respect to the effect of emotions on human memory. A body of research shows that emotions enhance memory in tone, while other claims that emotions enhance central information at the cost of peripheral details [20]. Considering these claims, it would be interesting to find answers to the following Research Questions (RQs) in the context of children-robot interaction:

RQ1 - Which of the following has a better effect on the child's learning outcome of vocabulary; a robot displaying positive, negative or a neutral emotional expression and gestures?

RQ1a - How does a robots positive and negative emotions affect the learning of the child across the single session?

RQ1b - How does a robots positive and negative emotions affect the learning of the child during a long-term interaction?

We hypothesize that a robot reacting positively to the child's vocabulary learning outcome will positively influence children learning as compared to negatively or neutrally during both single session and long-term interaction session.

3.1. Interaction Scenario

We programmed the NAO robot to autonomously play the game with children and teach vocabulary to them, however, speech recognition was controlled via a Wizard of Oz (WoZ) setup. We implemented a program to reply to basic preconceived questions during introduction and gameplay phases learnt from our previous study [22]. A facilitator responded to participant's queries through the WoZ setup. The robot stayed quiet in case, where the child asked questions out of its scope.

Learning Outcome	**NAO's Behaviour on the Outcome**
correct word	**Positive Condition:** I am delighted to know you got this correct, I am so happy that in the <SESSION NO> session, you got it wrong but this time your answer is correct, Happy or Joy Gestures **Negative Condition:** Ok, this is fine that you have got it correct, You remembered it today, but it took you <NO OF ATTEMPS> to learn <WORD NAME>. **Neutral Condition:** This is fine. let's check the next one.
incorrect word	**Positive Condition:** It is alright, I am hopeful you can get it right next time. **Negative Condition:** I am so sad, that you didnt remember <WORD NAME>, I am feeling disappointed that you don't remember <WORD NAME>, we learnt it in the <SESSION NO> session, Sad Gesture **Neutral Condition:** this is Ok, you can get next one right.
total correct words in the test	**Positive Condition:** I am so fond of you, you are performing consistently well, I am so proud of you, you are performing consistently well in all you tests, Clapping, bow down Gesture **Negative Condition:** you are doing well, but I am still sad that you answered <NUMBER OF WORDS> incorrectly, Disappointment Gesture **Neutral Condition:** This is alright, I think, lets hope for the best next time.

Table 2: *Robot's varying behaviour on Children's learning outcome.*

Our interaction session had three main phases. In the first phase, NAO robot asked about the words to be taught during the game. In the first session, NAO asked about the words with an assumption that children didn't know the word. The NAO responded by saying "We will learn about the <WORD NAME>

Game Event / Learning State	Emotional State	Event Type	Information	NAO's Behaviour
Snake near 100	Happy Smiling Surprise	Negative	User's emotional state, No. of times a snake has been received near 100	First Session: A snake near 100, I can see you are feeling <USER EMOTION>, its looks you want to learn more. Other Sessions: You had a snake near 100 during <SESSION NO> session, you looked < USER EMOTION >.
Snake near 100	Angry Sad Fear	Negative	User's emotional state, No. of blocks away from 100	First Session: This is <SNAKE OCCURRENCE> time you are on a snake today, but you look <USER EMOTION> about it, let's learn another word. Other Sessions: You looked < USER EMOTION > in session <SESSION NO> on a snake near 100, Are you not enjoying vocabulary learning today.
Ladder near 100	Angry Sad Fear	Positive	Occurrence of ladder and number of levels skipped	First Session: Why are you looking <USER EMOTION>, you are moving towards 100. Other Sessions: You also had a ladder near 100 in the <SESSION NO> session, I am happy to see you progressing well.
Ladder near 100	Happy Smiling Surprise	Positive	Occurrence of ladder and number of levels skipped	First Sessions: You look <USER EMOTION>, It is good to see you are playing well. Other Session: You also had a ladder near 100 in the <SESSION NO> session, you are extremely lucky.

Table 1: *Taxonomy for Emotional Event type and type of Information stored based on the Emotional Events*

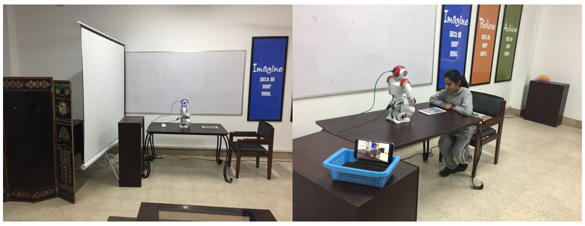

Figure 3: *Setup: A Child playing snakes and ladders with NAO.*

shortly". From the second to the fourth session, the robot provided positive, negative and neutral feedback through combining gestures and dialogue on the learning performance of each child. In the second phase, NAO robot played snakes and ladders game with the child. During the game, the child was taught six different words in each session. We coded a "fixed/predetermined" pattern of turns for both child and robot during the game for every session for all the participants. The afore-explained model was applied during the gameplay to create the memory of emotional events during the game. This memory was utilised after the first session. In the last phase, NAO tested about the words taught during gameplay. The NAO robot provided positive, negative, and neutral feedback that other than the first phase. In table 2, we present examples of the NAO robot's behaviour for the three categories of feedback during the pre- and post-test on children's learning performances.

3.2. Setup and Materials

We were assigned a quiet room at the school that was divided into two parts with a divider as shown in figure 3. On the left side, one of the researchers was controlling the speech recognition capabilities of the robot. On the right side, the child interacted with the NAO robot placed on the table along with a Samsung tablet. We used the NAO robot designed and developed by Aldebaran robotics. It is a humanoid robot measuring 58 cm in height with 25 degrees of freedom.

We used 24 vocabulary words from the Robot Interaction Language (ROILA) taken from the first two chapters of the book on ROILA [14].

3.3. Participants

We conducted our between-subject study with 24 children (12-males, 12 females) aged between 10-12 at a school. Each group comprised 8 children with equal ratio for the gender. None of the participants had previously interacted with a robot.

3.4. Procedure

Our study was setup as a long-term between-subject evaluation that lasted for three weeks. The study was conducted individually with one child at a time. Each child played the snakes and ladders game with the NAO robot 4 times for 4 days (one session per day), for a total of 96 sessions (24 child * 4 sessions). Each group of children played the game on a tablet in one of the three conditions (robot's displaying positive, negative, neutral emotional expression on child's learning outcome) for three school weeks. We conducted our sessions on the 1st, 5th, 10th and 14th day respectively. Each session lasted for approximately 24 minutes and had five steps: 1) a 4-minute pre-test, 2) a 10-minute game playing session, 3) a 4-minute break, and 4) a 4-minute post-test. The facilitator used a stopwatch to maintain time consistency throughout the sessions.

Introduction: NAO introduced itself and communicated with the child through a high-level dialogue. The dialogue involved inquiring about their day and activities that they are performing today.

Pre-test: The robot initiated the session through asking about unknown words from the ROILA language. The robot asked about six words during the first session. The rationale for selecting 6 words in each session came from a pilot study conducted with 5 participants. In the following sessions, six new words were added to the test. Therefore, 6, 12, 18 and 24 words were tested for in the first, second, third and fourth session respectively. The pre-test was an auditory-visual word identifica-

Session Nr.	Mean	S.D.
1	4.7500	.98907
2	5.3750	.76967
3	5.7500	.60792
4	5.0000	.65938

Table 3: *Mean values of children's learning outcome during sessions*

tion task [23]. The visual used in the test to represent a word was identical to the one used in the game playing sessions.

Game Play: Each child played the snakes and ladders game following a pre-defined pattern of dice outcomes for four times. Each child faced a snake inside the game six different times. On each snake, a new word was taught to the child. Therefore, in each session, six new words were taught to the participants.

Post-Test: After a 4-minute break,the child participated in the post-test to determine the accuracy of words learned during the session. The same procedure as pre-test was repeated in this phase, however, in the case of a mistake, the robot mentioned the correct answer during the feedback. The post-test was identical to the pre-test, containing the same words. We chose identical words for both pre and post-test to maintain the consistency of test results. All of the test results and mistakes were logged in the database.

3.5. Measurements

We measured the following Dependent variables: 1) total number of words learned during all the sessions (total number of words remembered in the last post-test on the last day), 2) immediate retention of new words during the session (total number of words remembered in the post-test of every session) and 3) retention of old words across sessions (total number of words remembered during the pre-test taught in the previous sessions).

4. Results

We conducted KolmogorovSmirnov test to ensure that the generated data was normally distributed before conducting Analysis of Variance (ANOVA). The results showed that the data was normally distributed for the learning outcomes of the children.

In order to check our model, we checked for the immediate retention of words learnt during each gameplay in all the sessions for all the participants. We conducted a repeated measure ANOVA with the $session$ as the within-subjects factor with four levels using immediate retention of words learnt per session as a Dependent Variable (DV). Results showed that there was a significant effect of session (F(2, 21) = 6.65 $p < 0.01$) on children's vocabulary learning. We executed Bonferroni posthoc to further examine the effect of learning outcome within sessions. We witnessed a significant effect when comparing learning outcome during the first and third session ($p < 0.01$) and during the third and fourth session ($p < 0.01$). The mean values of the learning outcome for all the sessions are shown in table 3. We found that children were able to learn words from the NAO robot capable of generating responses through following our emotion and memory model.

We conducted a one-way between-subject ANOVA with robot emotional feedback type as the independent variable (IV) and using a total number of words learned during all the session as a DV. The purpose of our analysis was to measure the overall effect of the type of emotional feedback on child's overall vocabulary learning. Our results show that there was a significant effect (F(2, 19) = 5.7 $p < 0.02$) of the type of robot's emotional feedback on the child's learning outcome. The mean retention rate across conditions were as follow: Condition 1) **M:** 22.25, **S.D.:** 1.03510, Condition 2) **M:** 20.25, **S.D.:** 1.16496 and Condition 3) **M:** 20.50, **S.D.:** 1.60357. We performed a Bonferroni posthoc check to further examine this significant difference. We found that robot with a positive emotional feedback had the better effect on child's overall learning outcome as compared with negative ($p < 0.04$) and neutral feedback ($p < 0.02$). The neutral feedback was significantly preferred ($p < 0.04$) as the second choice and negative was the last choice.

To check the effect of robot's feedback on children learning of words across sessions, we conducted a repeated measure ANOVA with the session as the within-subjects factor with $three$ levels and type of emotional feedback as the between-subject factor using retention of words learnt during the first session across sessions. We found a significant effect ($p < 0.04$) of robot's emotional feedback on the retention of words learnt during the first session. The Bonferroni posthoc check showed that positive emotional feedback from the robot has a better effect on child's learning ($p < 0.04$) as compared with negative feedback. We also conducted repeated measure ANOVA with the session as the within-subjects factor with two levels and type of emotional feedback as the between-subject factor using retention of old words learnt during the second session in the third and fourth session. We found a significant positive effect ($p < 0.02$) of robot's emotional feedback on the retention of words learnt during the first sessions. The Bonferroni posthoc check showed that robot with a positive emotional feedback has the significant effect on child's overall learning outcome ($p < 0.03$) as compared to the negative feedback. For the fourth session words, we conducted a one-way between-subject analysis of variance (ANOVA) with robot behaviour type as the IV and using a total number of words retained during the fourth session as DV. Our results showed that there was a significant effect ($p < 0.001$) of robot's emotional expression on the child's learning outcome. The Bonferroni posthoc check showed that robot with both positive and negative emotional feedback has the significant effect ($p < 0.03$) on child's overall learning outcome as compared to the neutral feedback.

5. Discussions and Conclusions

Our emotion and memory model for the NAO robot resulted in positive findings as the immediate retention of vocabulary for all the children was satisfactory during all the sessions. We witnessed a difference in children's learning outcomes within sessions. We conjecture that the reason might be due to NAO's novelty effect during the first session as children might be excited to see the robot. Similarly, a slight decline in the fourth session on the learning performance can also be grounded in the varying levels of children's interests in the interaction. As it is known that children learning can be affected due to the fall in their interest [7].

We also found that the positive emotional feedback provided by the robot on a child's learning performance did have a significant effect on the short-term and long-term children's vocabulary learning. Our hypothesis was accepted as children were able to retain the most number of words during robot's positive emotional feedback condition. In addition, the neutral response on child's feedback was also found to have an influence on child's learning. Moreover, the negative emotions were least regarded in terms of children's long-term learning performance. We conjecture that when the robot positively empathises with the child during the interaction, it creates a positive effect on the child's development in general. In the past, the role of emphatic robotic behaviour has been appreciated during a playful interaction as it was able to sustain children's interest in a long-term child-robot interaction [8]. We also speculate that the reason children were able to better retain words during positive emotional feedback is due to the positive emotional state of the children in response to the robot's reaction. It is shown in literature that humans store different kinds of information during different emotional states [20]. In addition, positive emotional feedback is conducive to feeling confident and successful in the learning process [24]. It would have facilitated enhanced learning. Therefore, we believe that a positive reaction of a robot created a positive emotional state of the child. It in return, made the child perform better as compared with the negative or neutral reaction of the robot. One of our findings also showed that from the third to fourth session, the negative feedback was preferred over neutral feedback. We speculate that although positive feedback is highly desirable, but negative feedback is also needed during the learning process. As it is shown in previous literature that negative feedback in terms of criticism may positively encourage student engagement and attention on learning task [25]. Therefore, negative feedback or directive critique can be useful in certain situations.

In general, our findings highlight the need for robotic tutoring systems where the robot takes a positive role and appreciate children. In addition, we also found a little evidence of using negative feedback such as displaying sadness during feedback may also lead towards improved learning. In Summary, it indicates towards implementing positive adaptive roles for the robotic tutor. The role can be a helper, a friend or a buddy that may also provide criticism at times during cHRI as also indicated by the teachers in one of the past studies [26].

5.1. Limitations and Future Work

One may argue on the number of sessions when it comes to being categorised as "long-term". However, our selection of a number of sessions is based on the findings of our previous study [22], where children's novelty factor diminished from the third session.

We understand that the number of repetitions of words as a part of feedback was not constant between the conditions. It might have created a potential confound in the results. In future, we will try to take it into account as a part of our data analysis.

We also understand that some participants would speak more than one language and/or have a greater level of capacity for learning another language. Our study did not select children on this basis.

In the future, we intend to test the generality of our model with multiple learning tasks (maths, science) during different games. We also intend to utilise our findings based on the feedback in future studies.

6. References

[1] P. B. Lawhead, M. E. Duncan, C. G. Bland, M. Goldweber, M. Schep, D. J. Barnes, and R. G. Hollingsworth, "A road map for teaching introductory programming using lego© mindstorms robots," in *ACM SIGCSE Bulletin*, vol. 35, no. 2. ACM, 2002, pp. 191–201.

[2] O. Mubin, C. J. Stevens, S. Shahid, A. Al Mahmud, and J.-J. Dong, "A review of the applicability of robots in education," *Journal of Technology in Education and Learning*, vol. 1, pp. 209–0015, 2013.

[3] D. Szafir and B. Mutlu, "Pay attention!: designing adaptive agents that monitor and improve user engagement," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 2012, pp. 11–20.

[4] J. B. Janssen, C. C. van der Wal, M. A. Neerincx, and R. Looije, "Motivating children to learn arithmetic with an adaptive robot game," in *International Conference on Social Robotics*. Springer, 2011, pp. 153–162.

[5] I. Leite, C. Martinho, and A. Paiva, "Social robots for long-term interaction: a survey," *International Journal of Social Robotics*, vol. 5, no. 2, pp. 291–308, 2013.

[6] A. Tapus, M. J. Mataric, and B. Scassellati, "Socially assistive robotics [grand challenges of robotics]," *IEEE Robotics & Automation Magazine*, vol. 14, no. 1, pp. 35–42, 2007.

[7] T. Kanda, T. Hirano, D. Eaton, and H. Ishiguro, "Interactive robots as social partners and peer tutors for children: A field trial," *Human-computer interaction*, vol. 19, no. 1, pp. 61–84, 2004.

[8] I. Leite, G. Castellano, A. Pereira, C. Martinho, and A. Paiva, "Empathic robots for long-term interaction," *International Journal of Social Robotics*, vol. 6, no. 3, pp. 329–341, 2014.

[9] M. I. Ahmad, O. Mubin, and J. Orlando, "Effect of different adaptations by a robot on children's long-term engagement: An exploratory study," in *Proceedings of the 13th International Conference on Advances in Computer Entertainment Technology*. ACM, 2016, p. 31.

[10] T. Belpaeme, P. E. Baxter, R. Read, R. Wood, H. Cuayáhuitl, B. Kiefer, S. Racioppa, I. Kruijff-Korbayová, G. Athanasopoulos, V. Enescu *et al.*, "Multimodal child-robot interaction: Building social bonds," *Journal of Human-Robot Interaction*, vol. 1, no. 2, pp. 33–53, 2012.

[11] G. Trafton, L. Hiatt, A. Harrison, F. Tamborello, S. Khemlani, and A. Schultz, "Act-r/e: An embodied cognitive architecture for human-robot interaction," *Journal of Human-Robot Interaction*, vol. 2, no. 1, pp. 30–55, 2013.

[12] I. Leite, "Long-term interactions with empathic social robots," Ph.D. dissertation, Instituto Superior Tcnico, Universidade de Lisboa, 2013.

[13] J. E. LeDoux, "Emotional memory," *Scholarpedia*, vol. 2, no. 7, p. 1806, 2007.

[14] O. Mubin, "Roila: Robot interaction language," Ph.D. dissertation, Eindhoven University of Technology, The Netherlands, 2011.

[15] C. Moskovsky and F. Alrabai, "Intrinsic motivation in saudi learners of english as a foreign language," *The Open Applied Linguistics Journal*, vol. 2, no. 1, 2009.

[16] D. Rinaldi and A. Ispita, "Teaching vocabulary by using capsule vocabularystrategy at fourth grade of elementary school," *e-Jurnal Mahasiswa Prodi Pend Bahasa Inggris 2012*, vol. 1, no. 4, 2012.

[17] L. S. Vygotsky, *Mind in society: The development of higher psychological processes*. Harvard university press, 1980.

[18] L. P. Rieber, "Seriously considering play: Designing interactive learning environments based on the blending of microworlds, simulations, and games," *Educational technology research and development*, vol. 44, no. 2, pp. 43–58, 1996.

[19] G. H. Bower, "How might emotions affect learning," *The handbook of emotion and memory: Research and theory*, vol. 3, p. 31, 1992.

[20] L. J. Levine and D. A. Pizarro, "Emotion and memory research: A grumpy overview," *Social Cognition*, vol. 22, no. 5: Special issue, pp. 530–554, 2004.

[21] M. Ahmad, O. Mubin, and P. Escudero, "Using adaptive mobile agents in games based scenarios to facilitate foreign language word learning," in *Proceedings of the 3rd International Conference on Human-Agent Interaction*. ACM, 2015, pp. 255–257.

[22] M. I. Ahmad, O. Mubin, and J. Orlando, “Adaptive social robot for sustaining social engagement during long-term children–robot interaction,” *International Journal of Human–Computer Interaction*, pp. 1–20, 2017.

[23] J. A. McCullough, L. A. Cunningham, and R. H. Wilson, “Auditory/visual word-identification test materials: Computer application with children,” *Journal of the American Academy of Audiology*, vol. 3, no. 3, pp. 208–214, 1992.

[24] J. Hattie and H. Timperley, “The power of feedback,” *Review of educational research*, vol. 77, no. 1, pp. 81–112, 2007.

[25] K. R. Wentzel, “Are effective teachers like good parents? teaching styles and student adjustment in early adolescence,” *Child development*, vol. 73, no. 1, pp. 287–301, 2002.

[26] M. I. Ahmad, O. Mubin, and J. Orlando, “Understanding behaviours and roles for social and adaptive robots in education: teacher's perspective,” in *Proceedings of the Fourth International Conference on Human Agent Interaction*. ACM, 2016, pp. 297–304.